基于智能服务的互联网直播信息管理系统的设计与实践

王可佳
（新华社通信技术局，北京 100803）

摘　要：随着信息技术及移动互联网技术发展的日新月异，当今舆论生态、媒体格局、传播方式均发生了深刻变化。尤其是媒体融合发展趋势下的视频直播业态，使传统广电模式面临着巨大挑战，实现全面 IP 化已是大势所趋。在这一背景下，如何更好地把握业务资源现状，以应对不断膨胀的业务规模及 4K/8K、VR 等超高清报道需求，实现高效的业务管理及信息共享，是必须直面的课题。同时，日趋成熟的智能化服务可以实现对直播内容有效信息的进一步提取及深度价值挖掘，这也为辅助全流程追踪、拓展统计分析维度、便捷实现分发共享、深度赋能融媒体生产协作等提供更多助力。本文将结合上述思路，围绕如何设计、构建基于智能服务的互联网直播信息管理系统，以及如何实现直播资源内容深度挖掘、共享等问题，展开进一步论述。

关键词：移动互联网技术；智能化服务；直播信息管理；数据采集；数据挖掘　　**中图分类号：**TN93　　**文献标识码：**A
文章编号：1671-0134（2021）09-135-03　　DOI：10.19483/j.cnki.11-4653/n.2021.09.043
本文著录格式：王可佳 . 基于智能服务的互联网直播信息管理系统的设计与实践 [J]. 中国传媒科技，2021（09）：135-136，67.

导语

近年来，新华社已经建成 IP 化的高清视频直播传输、调度网络，支持 64 路以上视频信号的同时接入及多渠道分发，有力支撑了新华社视频生产制作及直播报道业务的迅猛发展。然而，在媒体融合发展的大势之下，视频直播报道的业务规模不断膨胀，4K/8K、VR 等超高清、高带宽的报道需求持续增多，有限资源很难满足无限增长的业务需要。为了更直观地把握直播报道业务现状及手中的报道资源，更合理的规划报道安排及技术实施方案，需要对直播时间、内容、传输通路、人员、演播室、编解码设备、传输背包、摄像器材等资源及直播业务全流程各环节的状态进行动态管理和实时追踪；同时，新华社视频报道现在每月平均完成近百场次直播，这些内容传播效果如何，有无再利用、再生产价值，如何更高效地实现内容共享等问题，也都是报道业务的核心关切内容。希望在付出较小人工代价成本的前提下，进一步深入挖掘更多具有参考价值的数据信息，为未来组织报道提供决策依据。

1. 设计原则

围绕上述核心需求，新华社设计了一个基于智能服务的互联网直播信息管理系统，并在 2020 年年底完成该系统的基本功能设计与第一阶段的自主开发工作。这一平台能为直播报道全流程业务管理提供了更有力的支撑，也为后续代码迭代及功能升级奠定了坚实基础。这一系统采用如下设计原则。

1.1　移动性

在实现直播全流程业务管理功能、为 Web 端访问提供更直观交互服务接口的基础上，充分利用移动互联网和 5G 技术的发展优势，为移动端提供更好的界面呈现效果及互动操作体验。同时，构建运营级的高可用开源流媒体服务组件并引入具有跨平台、先进性的轻量级 HTML5 流媒体播放插件，为移动端直播流监看等场景提供更清晰、更流畅的播放体验。

1.2　高效性

除实时向多渠道媒体用户分发直播信号外，直播保障团队成员也需要深度参与直播报道工作，对直播流及传输通路的状态进行监控，对演播室、设备、人员等报道资源进行统筹调配；除此之外，大数据监控平台需要日常展示直播信号，报道调度等系统需要在重大报道时展示多路实时回传画面。这要求系统能够动态获取、实时呈现资源占用情况，并引入更加高效的流媒体传输协议与数据交换机制，服务于更多用户及上下游展示端对直播信号的共享需要。

1.3　智能化

构建智能服务模块，自动生成每日直播预告及直播简报，对直播报道内容进行深度挖掘，辅助完成多渠道数据采集及统计数据汇总，降低全流程业务管理的人工成本。

1.4　易联通

提供更好的联通拓展性，利用通用服务接口灵活对接直播调度系统、视音频内容汇聚分发平台 [1] 等多个上下游系统，建立高效局域网内传输机制，便捷实现视频内容与元数据的分发共享。

1.5　易维护

系统稳定可靠，为应用软件及中间件等采用虚拟化部署方式，完善数据备份及快照机制；对应用程序运行引入服务化启停及日志监控机制，做到易管理、易维护、易恢复。

2. 总体架构

根据上述思路，基于智能服务的互联网直播信息管理系统架构的设计如图 1 所示。



图 1　基于智能服务的互联网直播信息管理系统架构

系统以虚拟化、组件化、智能化的技术路线为主要出发点，在虚拟化平台上搭建 Web 服务器、智能化引擎及其他必备的基础资源，从而支撑上层各组件、模块、服务的实现：搭建 Http Server，选用以 Python Django 为基础的开源 Web 应用框架，实现模板化、松耦合、易扩展的网页功能及服务封装；系统引入对象存储 +Mysql 关系型数据库 +Redis 内存数据库等中间件服务构成的多级存储结构，从根本上提升全流程数据的存储管理效率，为数据采集及统计分析等业务功能提供底层支持；系统还通过部署消息队列服务集群，实现业务、统计数据的分发及流程实时调度，并通过搭建 TensorFlow、Paddle Paddle 等智能化引擎，为实现智能化服务的核心模块功能及用户服务层的应用逻辑等构建基础环境。

在基础资源之上，系统也设计了基础组件层和核心模块层：引入组件化数据采集服务，多渠道追踪新华社内外用户对直播报道的采用及传播情况；搭建智能化服务模块，实现 OCR 图片文字识别、视频内容目标检测等功能，深度提取并挖掘直播报道内容中的视频与文本信息；设计通用数据交换服务，为上下游系统间的直播流、视频文件及元数据交换等提供对接标准及共享支持；构建高可用的流媒体服务组件，引入低带宽、低延时、适用于局域网场景的视频流传输协议，为直播流实时监看、多端展示等提升效率。

同时，系统设计用户服务层，面向用户提供较为完善的直播信息管理服务：实现直播信息管理业务，支持直播流及元数据信息管理、直播日历展示等功能；设计直播设备管理服务，实现视频编解码器、传输背包、摄像机、VR 摄影机等直播报道设备的查询、预约、入库、借出、归还、审核等管理操作；设计演播室预约管理服务，支持对直播业务中非常关键的演播室、传输光纤等资源实现查询、预约、分配、审核等管理功能；设计直播人员管理服务，支持直播业务制作及技术团队人员管理功能，提供人员权限及排班管理、直播预告、短信通知、直播简报生成、直播流在线监看等功能。

3. 核心功能

系统以智能化服务、数据采集服务及高可用流媒体服务等组件化服务为基础，实现直播业务管理、直播采用分析及直播信号监看共享等核心功能。

3.1　直播业务管理

系统实现直播信息管理服务，支持直播流及元数据信息管理功能。除实现相关信息的人工管理外，系统还充分利用智能化服务，达到简化用户操作及进一步强化内容挖掘的目标。系统支持直播信息一键式录入功能，利用智能 OCR 识别（图片文字识别）服务自动实现业务申请单中的文字识别及电子化转换功能，利用智能自然语言处理服务实现申请单中时间、地点、人员等关键信息提取，进而完成对直播信息的结构化归档，为直播全流程业务管理及内容价值的深度挖掘提供有力支撑。此外，系统还支持利用第三方的智能语音转写、智能视频分类及图像目标检测服务实现视频内容中有效信息的提取。

3.2　直播采用分析

系统通过组件化数据采集服务，定时、多渠道追踪新华社内外用户的直播报道采用信息及影响力数据。系统采集互联网用户平台中新华社直播报道基本数据及点击量、播放量、用户评论、点赞量等传播效果指标信息。采集后的数据使用内存数据库集群进行查重，将数据写入分布式日志系统实现异步归档、分发。

系统以上述采集到的影响力数据为基础，向用户提供分析报表，总结每月直播报道数量及地域分布情况，并借助智能化服务中的内容相关度检测、评论倾向分析等功能，建立传播影响力数据模型，对传播效果实行量化分析，生成报表，为未来报道策划提供决策依据。

3.3　直播监看共享

搭建开源流媒体服务组件，除支持 HTTP、RTMP、HLS、WEB RTC 等主流视频直播 / 点播协议外，引入具有低带宽、高可靠、低延时等特性的 SRT、NDI 协议，打造更高效并适用于局域网场景的直播 / 点播视频分发框架。如图 2 所示。



图 2　直播信息管理系统的直播流监看页面

（下转第67页）